

On the Security of Interferometric Quantum Key Distribution

Ran Gelles^{1,*} and Tal Mor²

1. *Computer Science Department,*
UCLA, Los Angeles, USA,
 gelles@cs.ucla.edu

2. *Computer Science Department,*
Technion, Haifa, ISRAEL,
 talmo@cs.technion.ac.il

Abstract

Photonic quantum key distribution is commonly implemented using interferometers, devices that inherently cause the addition of vacuum ancillas, thus enlarging the quantum space in use. This enlargement sometimes exposes the implemented protocol to new kinds of attacks that have not yet been analyzed.

We consider several quantum key distribution implementations that use interferometers, and analyze the enlargement of the quantum space caused by the interferometers. While we prove that some interferometric implementations are robust (against simple attacks), we also show that several other implementations used in QKD experiments *are totally insecure*.

This result is somewhat surprising since although we assume ideal devices and an underlying protocol which is proven secure (e.g., the Bennett-Brassard QKD), the realization is insecure. Our novel attack demonstrates the risks of using practical realizations without performing an extensive security analysis regarding the specific setup in use.

* Part of this work done while at Technion, Israel.

I. INTRODUCTION

Quantum Key Distribution (QKD) is a cryptographic protocol for expanding a pre-shared secret between two users (Alice and Bob) by transferring quantum systems. Once an adversary (Eve) tries to acquire information about a transferred quantum system, she inevitably disturbs it in a way that can be detected by the legitimate users, and causes the abortion of the protocol; this principle is known as “Information Vs. Disturbance” [1–6].

The first and most popular QKD protocol is the BB84 protocol [7], in which Alice sends qubits to Bob using two conjugate bases. In the real world, qubits are implemented via various methods. A very common QKD implementation is the *phase-encoded, time-multiplexed* scheme: A pulse that contains a single photon is sent in a superposition of two possible times, so that the encoded bit is the phase difference between these superpositions as initially suggested by Bennett [8] and implemented by Townsend and others [9–11]. In order to produce and measure such superpositioned pulses, it is common to use an *interferometer* (see Section IV A). In addition to the basic (phase-encoded, time-multiplexed) setup, interferometers are also used in more complex QKD setups. For instance, an interferometer is used in the implementation of *Differential Phase Shift QKD (DPS-QKD)* [12–14], which generalizes the time-multiplexing scheme by encoding each bit as a phase shift of three superpositioned pulses (instead of two). Another variant which uses interferometer is the *Plug & Play* protocol used in many experiments [15, 16] and commercial products [17, 18], in which the signal is generated by Bob, sent over to Alice who modulates its phase, then sent back to be measured by Bob.

In this paper we analyze interferometric based QKD schemes. Specifically, we discuss different ways to implement BB84 using interferometers. We also discuss implementations of more general schemes, such as the six-state QKD protocol. Once a protocol is implemented via photons and interferometers, the implementation differs from the ideal protocol (that uses abstract qubits) since the “ideal world” two-dimensional qubit space is replaced with a “real world” larger quantum space. This is due to two reasons: first, interferometers inherently introduce a higher-dimension space; and second, having pulses with zero photons, or more than one photon, implies a higher dimension as well. Here we focus on the first space enlargement.

The usage of an enlarged quantum space requires a more careful security analysis of such implementations. Since Eve controls this large space (or parts of it), rather than the ideal qubit space, she can perform a much stronger attack than on the theoretical protocol. In this paper we design a novel type of attack, the *reversed-space attack*, based on considering this large space.

As proving the security of a scheme is usually difficult, in this work we consider *robustness*, the ability to identify attacks on the protocol. Although robustness is a weaker property than security, showing robustness might be a first step towards a full security analysis. On the positive side, we show that many of the interferometric implementations are indeed robust (against a limited adversary), which might hint at the security of these implementations. On the other hand, we demonstrate a reversed-space attack on several BB84 implementations used in recent experiments [19–22] proving them to be insecure. Other realizations (e.g. [12, 23–25]), extend the above variants and use an even larger Hilbert space. The security of such extensions should be analyzed as well, potentially using the tools we present here.

This work joins a line of research that examines the security of QKD implementations. Although BB84 has been proven secure against the most powerful attacks [6, 26–29], these proofs do not apply to realistic variants, and specific attacks were presented to exploit limitations of specific implementations (e.g., [30]). Several security analyses have been published for special cases [31, 32]: e.g., a specific protocol variant (DPS-QKD, Plug&Play, etc.), or a specific eavesdropping method. In addition, recent analyses have considered the security of protocols realized using imperfect

equipment, such as faulty sources and detectors [33–35]. Still, a general framework that considers such realistic QKD protocols, *as well as* attacks on such protocols, is still missing.

A. The BB84 Protocol

We mainly focus on implementations of the BB84 protocol [7]. In BB84 Alice uses two conjugate bases (say, z and x) to encode each of her bits as one of the states $|0_z\rangle$, $|1_z\rangle$, $|0_x\rangle = \frac{1}{\sqrt{2}}(|0_z\rangle + |1_z\rangle)$, or $|1_x\rangle = \frac{1}{\sqrt{2}}(|0_z\rangle - |1_z\rangle)$. Let H^A be the quantum space Alice holds (here Alice is ideal, therefore $H^A = H_2$).

The security of the protocol is based on limiting the error-rate measured by Bob, under the assumption that when no eavesdropping has occurred, Bob should perfectly retrieve the bits encoded by Alice. See more details in, e.g., [4–6, 27].

Note that in the theoretical (ideal) protocol, Alice and Bob use the same two-dimensional space H_2 . However, in a non-ideal world, the spaces H^A and H^B , held by Alice and Bob respectively, might be larger. An important demonstration of such a case is given in [30], where a realistic photonic source is analyzed, such that the cases of zero photons and two photons are added, and as a result $\dim H^A = 6$. Surprisingly, when interferometers are used, even in the case where $\dim H^A = 2$ and all the devices are ideal, Bob measures six orthogonal states that are correlated to the pulse sent by Alice. If Bob actually measures these 6 states, his measured-space becomes much larger than H_2 , specifically, $\dim H^B = 6$.

B. Eavesdropping

An eavesdropper can perform various kinds of attacks, however we focus on individual-particle attacks. The most simple attack Eve can compose is a basic *measure-resend* attack, in which Eve measures the qubit sent by Alice and obtains a classical outcome. Then, Eve sends Bob a different qubit, determined by the classical outcome she measured. A much stronger individual-particle attack can be done by attaching a separate auxiliary particle (the ancilla $|0\rangle_{\tilde{E}}$) to each one of the qubits sent by Alice, and performing a unitary transformation $\mathcal{U}_{\tilde{E}}$ on each qubit along with its ancilla, possibly entangling them. To be more accurate, on each qubit sent by Alice, Eve performs a unitary transformation

$$|0\rangle_{\tilde{E}}|i\rangle_A \xrightarrow{\mathcal{U}_{\tilde{E}}} \sum_j \epsilon_{i,j} |\tilde{E}_{i,j}\rangle_{\tilde{E}} |j\rangle_A, \quad (1)$$

where i (and j) are vectors in the computation basis. Eve’s ancillas (denoted as the subsystem \tilde{E}) are kept for a later measurement, performed *after* Eve learns the basis used for each qubit. For attacking an ideal BB84 scheme, it is sufficient to have a small-dimension ancilla, namely $\dim H^{\tilde{E}} \leq 4$, due to Davies’ theorem [36].

C. Robustness

The criterion of *robustness* is often used in security analyses of QKD protocols (e.g., see robustness analysis of the SARG protocol [37], the BBM protocol [38] and the classical-Bob protocol [39]). We follow the robustness definition of [39] to analyze interferometric QKD implementations.

Definition 1. A protocol is said to be **completely robust** if nonzero information acquired by Eve implies nonzero probability that the legitimate participants find errors on the bits tested by the

protocol. A protocol is said to be **completely nonrobust** if Eve can acquire the entire information transmitted in the protocol (namely, the entire information string), without inducing any errors on the bits tested by the protocol.

Another closely-related definition is of *partial robustness* [39]: “A protocol is said to be partly robust if Eve can acquire some limited information on the information string without inducing any error on the bits tested by the protocol.” Partly-robust protocols could still be secure, yet completely nonrobust protocols are automatically proven insecure. As an illustrative example [30], BB84 is fully robust if Alice and Bob use qubits, but it is not completely robust if instead of (ideal) qubits they send pulses, and those pulses sometimes contain more than one photon.

Proving robustness of ideal protocols against any attack is easier than proving its security. For complicated protocols and for practical implementations it is common that robustness is proven first, and the security proof is left for future work.

D. Model and Assumptions

The focus of this paper is QKD implementations based on single *photons* as the quantum carriers. In order to describe a qubit using a single photon, one needs to define two possible orthogonal states. These orthogonal states (called *modes*) are commonly either an intrinsic property (e.g. the polarization of the photon, $|\uparrow\rangle$ and $|\leftrightarrow\rangle$), or a spatial separation (e.g., $|t_0\rangle$ and $|t_1\rangle$ for different times t_0, t_1). A very convenient way to describe photonic qubits is the *Fock Space* notations (See Appendix A). Although our analysis is done using Fock notations, we stick in this extended abstract to the standard notations, and refer the reader to the appendix for full details.

For simplicity, we assume throughout the paper that Alice’s operations are ideal, namely she always succeeds in generating a qubit in the exact desired state. Under this assumption, it is easier to see the novelty and importance of the attack that we suggest and analyze here.

We restrict the adversary to sending only pulses with a single photon, a *single-photon-limited Eve* (In Appendix C, we extend the robustness proof to an adversary which is limited to sending 2 photons). Nevertheless, the adversary is capable to receive, hold and manipulate quantum systems of higher dimensions. Moreover, all our robustness proofs are against individual-particle attacks.

Finally, we discuss the way losses are treated and differentiated from errors, since it can sometimes influence the robustness (and security) analysis, as we now explain. Security proofs (e.g., [6, 26, 27]) determine the maximal error rate (attributed to Eve’s attack) that keeps Eve’s knowledge negligible. If Bob considers each loss as a random bit obtained from Alice, he adds an error with probability half, and thus increases the error-rate. In reality, the loss-rate is too high (commonly, 90% or even 99%) for considering each loss as half-an-error, since the resulting error-rate will exceed the threshold, and the protocol will always be aborted. Allowing losses without defining a loss-rate threshold might allow Eve to perform useful attacks that result in losses yet no errors, and thus cannot be detected [30, 40]. In such cases, one might be able to define a loss-rate threshold such that for a high loss-rate the protocol is completely nonrobust, while for a low loss-rate the protocol is partly robust, and might yield a secure final key.

II. ATTACKS IN AN ENLARGED SPACE

We now adapt the standard security analysis to the case where Alice is ideal and Bob measures a larger space (for instance, Bob uses interferometer for his measurements). As a first step we define the set of Eve’s attacks on that large space that cannot be identified by the legitimate users,

that is, attacks that cause no errors. Later we consider the maximal amount of information Eve may obtain by performing such an attack.

A. Formulating Eve's Attack

Assume that Alice is ideal, and denote the basis states of her system by $|i\rangle_A$ with $i \in \{0, 1\}$. Eve adds an ancilla in the state $|0\rangle_{\tilde{E}}$ and performs her attack on the joint system A and \tilde{E} as described by Equation (1), using a unitary transformation $\mathcal{U}_{\tilde{E}}$. Eve continues her attack by sending the subsystem A to Bob.

However, Eve can use an enlarged system (e.g. by adding photons or modes). Eve has incentive to send states beyond the qubit of Alice if these influence Bob's measurement. A more general attack can be described as

$$|0\rangle_{\tilde{E}}|i\rangle_A \xrightarrow{\mathcal{U}_E} \sum_k \epsilon_{i,k} |E_{i,k}\rangle_E |k\rangle_P \quad (2)$$

where the subsystem P (rather than A) is sent over to Bob, and $|k\rangle_P$ are basis states of the system P . Obviously, $P = A$ is merely a special case, while in the more general case Eve might send Bob a system with different dimensions than A . The subsystem E , which remains in Eve's hands to be measured afterwards, can therefore differ from the subsystem \tilde{E} she initially had. Moreover, both can be of any dimension as long as the dimension of the entire system does not change, $H^{\tilde{E}} \otimes H^A = H^E \otimes H^P$. Let a general qubit sent by Alice be $|\psi\rangle_A = \sum_i \alpha_i |i\rangle$ with $\sum_i |\alpha_i|^2 = 1$, then due to linearity, the attack on that qubit is

$$\mathcal{U}_E \left(\sum_i \alpha_i |0\rangle_{\tilde{E}} |i\rangle_A \right) = \sum_{i,k} \alpha_i \epsilon_{i,k} |E_{i,k}\rangle_E |k\rangle_P. \quad (3)$$

B. Formulating Bob's Measurement

In order to perform a general measurement, Bob might manipulate the state sent by Alice. For instance, Bob might add an ancillary system B' , perform some unitary operation on the joint system and then perform a measurement of the joint system AB' .

We model Bob's measurement as (i) adding the ancilla¹ $|0\rangle_{B'}$; (ii) performing a unitary transformation \mathcal{U}_B on $|\psi\rangle_A |0\rangle_{B'}$; and then (iii) measuring the joint system in the computation basis. Note that \mathcal{U}_B changes according to the specific basis used by Bob. For the case of BB84, where Bob uses a separate setup for the x and the z basis, we get

$$|i\rangle_A |0\rangle_{B'} \xrightarrow{\mathcal{U}_{Bz}} \sum_j \beta_{i,j}^z |j\rangle_{AB'} \quad ; \quad |i\rangle_A |0\rangle_{B'} \xrightarrow{\mathcal{U}_{Bx}} \sum_j \beta_{i,j}^x |j\rangle_{AB'} . \quad (4)$$

The β 's are determined by the specific setup used by Bob, and the states $|j\rangle_{AB'}$ are Bob's basis states in the computation basis, that is, the set of states that span $H^A \otimes H^{B'}$.

As an illustrative example, assume that Bob adds no ancilla and his detectors are set to the z -basis. Therefore, for measuring the z basis Bob performs no transformation (i.e. $\mathcal{U}_{Bz} = I$, the identity matrix), while for measuring the x -basis he must perform the Hadamard transformation $\mathcal{U}_{Bx} = \mathcal{H} \triangleq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, so that $|0_x\rangle \xrightarrow{\mathcal{U}_{Bx}} |0_z\rangle$ and $|1_x\rangle \xrightarrow{\mathcal{U}_{Bx}} |1_z\rangle$. Since Bob adds no ancilla,

¹ Without loss of generality, we assume that Bob uses the same ancilla $|0\rangle_{B'}$ for all of his setups. This can always be justified, e.g., by using a sufficiently large ancilla, such that the different setups potentially use different subsystems of that ancilla.

$H^{AB'} = H^A$ in this case. Using the notations of Equation (4), we get $\mathcal{U}_{B_z}|0\rangle_A = |0\rangle_{AB'}$ thus $\beta_{0,0}^z = 1$ and $\beta_{0,1}^z = 0$; $\mathcal{U}_{B_z}|1\rangle_A = |1\rangle_{AB'}$ thus $\beta_{1,1}^z = 1$ and $\beta_{1,0}^z = 0$. We can write β as a matrix $\beta = (\beta_{i,j})_{i=0..1,j=0..1}$ which gives $\beta_{i,j}^z = I$, and in a similar way $\beta_{i,j}^x = \mathcal{H}$.

It should be noted that Bob's transformation \mathcal{U}_B is actually defined on Bob's ancilla and the space entering his lab, H^P (rather than H^A). Eve has no incentive to send Bob states that can *never* affect his measurement and we can assume Eve sends exactly H^P to Bob (any other system that cannot affect Bob is sent back to Eve and considered as part of her ancilla H^E). In this case, Bob performs a measurement of $B = PB'$, rather than of AB' .

While a fully powerful Eve knows the protocol space of Alice and Bob, and their equipment limitations, Alice and Bob themselves might not be aware of the fact that the system A is replaced by the enlarged system P . The above formulas (4) immediately generalize to this case, simply by replacing the subscript A by the subscript P . The general final state $|\Psi_{EB}\rangle$ (held by Bob and Eve), can be written as

$$|\Psi_{EB}\rangle = (I_E \otimes \mathcal{U}_B) \sum_{i,k} \alpha_i \epsilon_{i,k} |E_{i,k}\rangle_E |k\rangle_P |0\rangle_{B'} = \sum_{i,k,j} \alpha_i \epsilon_{i,k} \beta_{k,j} |E_{i,k}\rangle_E |j\rangle_B . \quad (5)$$

There is a great deal of importance regarding the way Bob interprets his measurement outcome. The states $|j\rangle_B$ can be classified into sets according to Bob's interpretation: some of these states indicate "Alice has sent the bit 0", others indicate "Alice has sent the bit 1". Let us denote these two sets by J_0 and J_1 , respectively. An error occurs when Alice sends a bit b , while Bob measures a state in J_{1-b} . Generally, for a specific transmission, we define by J_{error} the set of all states that imply an error, so in the example above $J_{\text{error}} = J_{1-b}$.

C. Attacks that Cause No Error

When considering real implementations, there may be some outcomes that are not interpreted as a valid outcome. These outcomes can be divided into two groups:

1. outcomes interpreted by Bob as a loss — a failed transmission that is not considered as an error, because they naturally occur even when no eavesdropper interferes (e.g. a vacuum state). These outcomes are denoted as the set J_{loss} .
2. invalid-erroneous outcomes J_{invalid} — outcomes that can never occur if the quantum system sent by Alice reaches Bob intact. It is Bob's choice of interpretation that determines whether a specific outcome is considered a loss or an invalid result. Generally speaking, when an invalid outcome increases the error rate, it is in J_{invalid} .

In order to analyze the robustness of QKD protocols, we consider attacks that cause no errors or invalid outcomes at Bob's end. Formally, for any $|j'\rangle$ in J_{error} or J_{invalid} , we require the overlap $\langle j' | \Psi_{BE} \rangle$ to be zero. Using Equation (5) we see that Eve's attack causes no errors if and only if $\langle j' | \sum_{i,k,j} \alpha_i \epsilon_{i,k} \beta_{k,j} |E_{i,k}\rangle_E |j\rangle_B = 0$, for any $j' \in J_{\text{error}} \cup J_{\text{invalid}}$.

Corollary 1. *For a given QKD implementation, Eve's attack \mathcal{U}_E causes no errors if and only if for every state $|\psi\rangle_A = \sum_i \alpha_i |i\rangle_A$ sent by Alice,*

$$\sum_{i,k} \alpha_i \epsilon_{i,k} \beta_{k,j} |E_{i,k}\rangle_E = 0 , \quad (6)$$

for any $j \in J_{\text{error}} \cup J_{\text{invalid}}$ (corresponding to the specific state $|\psi\rangle_A$).

Given a specific QKD implementation, the error rate is exclusively determined by the attack \mathcal{U}_E performed by Eve.

Definition 2. Let \mathbf{U}_{zero} be the set of attacks on a given protocol, that cause no errors (in all the possible setups of the protocol).

A scheme is robust if \mathbf{U}_{zero} only consists of attacks that give Eve no information about the key. For BB84 with bases z, x , \mathbf{U}_{zero} is determined by the intersection of the zero-error attacks for z -basis and x -basis.

III. THE POWER OF REVERSED-SPACE ATTACKS

One can apply the reversed transformation $(\mathcal{U}_B)^{-1} = \mathcal{U}_B^\dagger$ on each possible $|j\rangle_B$, in order to identify the space that influences Bob's outcome. We call an attack, designed according to this observation, a *reversed-space attack*. The term “reversed” here is borrowed from the “time reversal symmetry” of quantum theory. The symmetry of quantum mechanics to the exchange of the prepared (preselected) state and the measured (postselected) state was suggested by [41, 42] (and was already used in quantum cryptography as well, see the time-reversed EPR scheme [43]). To clarify this point, we stress that in Eve's attack described above, our assumption is that Eve sends Bob a quantum system with space H^P , instead of H^A . Eve has no incentive in sending systems of higher-dimension, thus she can limit H^P to be the space given by the reverse method (i.e., the space obtained by considering $(\mathcal{U}_B)^{-1}$). As a by product, this simplifies any security analysis, since there is no use in analyzing spaces of larger dimension than the reversed-space.

In the following sections we describe different BB84 implementations that use interferometers, and analyze their robustness via the reverse-space method, against an adversary that is limited to sending pulses with up to one photon.

IV. ANALYSIS OF PHASE-ENCODED INTERFEROMETRIC BB84

In this section we analyze a phase-encoded time-multiplexed QKD implementation [9], and show it is robust against a limited adversary.

A. Interferometric Implementation of the xy -BB84 Scheme

Consider a BB84 implementation which uses two time-separated modes (pulses). For every transmission, the first mode arrives to Bob's lab at time t'_0 , and the second mode at $t'_1 = t'_0 + \Delta T$. We denote these pulses as $|t'_0\rangle$ and $|t'_1\rangle$ respectively. The users use the x and y bases, so that an ideal Alice sends one of the following four states,

$$\begin{aligned} |0_x\rangle_A &\equiv (|t'_0\rangle + |t'_1\rangle) / \sqrt{2} & |0_y\rangle_A &\equiv (|t'_0\rangle + i|t'_1\rangle) / \sqrt{2} \\ |1_x\rangle_A &\equiv (|t'_0\rangle - |t'_1\rangle) / \sqrt{2} & |1_y\rangle_A &\equiv (|t'_0\rangle - i|t'_1\rangle) / \sqrt{2} . \end{aligned}$$

Bob measures the qubit using a Mach-Zender interferometer, which is a device composed of two beam splitters (BS) with one short path, one long path, and a controlled phase shifter P_ϕ , that is placed at the long arm of the interferometer. (see Appendix B for a full description of an interferometer, and analysis of its operation on single-photon modes). The length difference between the two arms is determined by ΔT : when the first pulse travels through the long arm, and the second through the short arm, they arrive together at the output. Due to that exact timing

of the pulses, each incoming qubit is transformed into a superposition of 6 possible modes: 3 time modes (t_0, t_1, t_2) at the straight (s) output arm of the interferometer, and 3 modes at the down (d) output arm; see Figure 1.

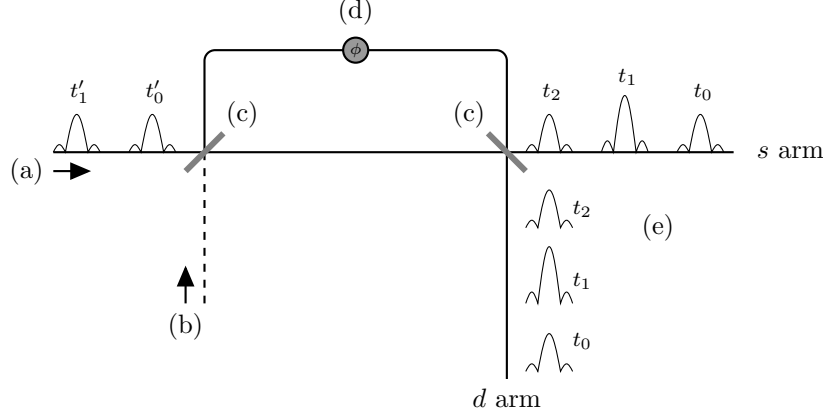


FIG. 1. A Mach-Zender interferometer. (a) An input qubit. The time-difference between the two incoming modes is identical to the difference between the two arms; (b) a vacuum state entering the second (blocked) arm; (c) beam-splitters; (d) phase shifter P_ϕ ; (e) six output modes.

For the sake of simplicity we denote these modes as $s_0, s_1, s_2, d_0, d_1, d_2$, and since we only consider pulses with zero or one photons, we can use the states, $|s_0\rangle, |d_0\rangle$, etc.², along with the vacuum state $|V\rangle$ (a pulse that contains no photons).

The interferometer's operation (See Equation (B2)) is defined by $|V\rangle \mapsto |V\rangle_B$ and

$$\begin{aligned} |t'_0\rangle &\mapsto (|s_0\rangle_B - e^{i\phi}|s_1\rangle_B + i|d_0\rangle_B + ie^{i\phi}|d_1\rangle_B)/2 \\ |t'_1\rangle &\mapsto (|s_1\rangle_B - e^{i\phi}|s_2\rangle_B + i|d_1\rangle_B + ie^{i\phi}|d_2\rangle_B)/2. \end{aligned} \quad (7)$$

Bob sets the phase ϕ according to the basis he wishes to measure: $\phi = 0$ for the x -basis and $\phi = \pi/2$ for the y -basis. When Alice's and Bob's bases match, the input qubit evolves in the interferometer as

$$\begin{aligned} |0_x\rangle_A &\xrightarrow{\phi=0} (|s_0\rangle_B - |s_2\rangle_B + i|d_0\rangle_B + 2i|d_1\rangle_B + i|d_2\rangle_B) / \sqrt{8} \\ |1_x\rangle_A &\xrightarrow{\phi=0} (|s_0\rangle_B - 2|s_1\rangle_B + |s_2\rangle_B + i|d_0\rangle_B - i|d_2\rangle_B) / \sqrt{8} \\ |0_y\rangle_A &\xrightarrow{\phi=\pi/2} (|s_0\rangle_B + |s_2\rangle_B + i|d_0\rangle_B - 2|d_1\rangle_B - i|d_2\rangle_B) / \sqrt{8} \\ |1_y\rangle_A &\xrightarrow{\phi=\pi/2} (|s_0\rangle_B - 2i|s_1\rangle_B - |s_2\rangle_B + i|d_0\rangle_B + i|d_2\rangle_B) / \sqrt{8} \end{aligned} \quad (8)$$

Bob opens his detectors at time t_1 at both the arms. A click at the “down” direction (i.e., measuring the state $|d_1\rangle$) means the bit-value 0, while a click at the “straight” direction ($|s_1\rangle$) means 1. The other modes are commonly considered as a loss (namely, they are not measured) since they do not reveal the value of the original qubit. The above implementation is commonly used for QKD experiments [44–46], and products [17, 18]. We denote this implementation scheme by xy -BB84.

Since measuring the other modes ($|s_0\rangle$, etc.) does not reveal the bit Alice has sent, measuring these modes can only help Bob in noticing some eavesdropping attacks. On the other hand, considering these modes complicates the security analysis since Eve might send superpositions of the time-modes $t'_2 = t'_1 + \Delta T$, and $t'_{-1} = t'_0 - \Delta T$, which will not result in $|V\rangle$.

² Using the Fock state notations and the description of interferometers in Appendix B, a basis state in Bob's space is given by $|n_{s_0}, n_{s_1}, n_{s_2}, n_{d_0}, n_{d_1}, n_{d_2}\rangle^F$, and we define $|100000\rangle^F \equiv |s_0\rangle$; $|010000\rangle^F \equiv |s_1\rangle$; $|001000\rangle^F \equiv |s_2\rangle$; $|000100\rangle^F \equiv |d_0\rangle$; $|000010\rangle^F \equiv |d_1\rangle$; $|000001\rangle^F \equiv |d_2\rangle$, and the vacuum state $|000000\rangle^F \equiv |V\rangle$.

B. Robustness Proof for a Single-Photon-Limited Eve

We now prove the *xy*-BB84 implementation to be completely robust against a single-photon limited Eve. We begin by defining the set \mathbf{U}_{zero} of attacks that induce no errors, according to Definition 2. Since we limit the parties to single-photon pulses, we only need to consider the subspace that contains the vacuum state and the single-photon states, i.e., the states $|V\rangle$, $|s_0\rangle$, $|s_1\rangle$, $|s_2\rangle$, $|d_0\rangle$, $|d_1\rangle$ and $|d_2\rangle$ defined above. The analysis must consider the space Bob actually measures, H^B , and the states (sent by Alice or Eve) that, after the interferometer, have a non-zero overlap with the modes measured by Bob. This idea is what differentiates our analysis from previous methods of analyzing (theoretical) protocols.

In *xy*-BB84, Bob only measures time-bin t_1 , thus we are interested in the subspace spanned by $\{|V\rangle, |s_1\rangle, |d_1\rangle\}$. By applying \mathcal{U}_B^{-1} on these three states, we get the states sent by Eve³ that influence Bob. Following Section II we refer to such an attack as the *reversed-space attack* on this specific scheme.

Theorem 2. *Assuming a single-photon-limited adversary, the *xy*-BB84 scheme is completely robust*

Proof. We describe the measurement Bob performs, for both bases, as $J_0 = \{|d_1\rangle\}$; $J_1 = \{|s_1\rangle\}$; $J_{\text{loss}} = I - J_0 - J_1 = \{|V\rangle\}$; and $J_{\text{invalid}} = \{\}$, where the set I represents the computation basis of the space measured by Bob, and the minus stands for set difference. Eve might send the photon at any desired time-bin, or not send a photon at all. However, the photon will affect Bob only in the case it is sent at time-bins t'_0 or t'_1 . Thus, Eve has no advantage in attacking a larger space than the one used by Alice, and H^P is spanned by $\{|V\rangle_A, |t'_0\rangle_A, |t'_1\rangle_A\}$. Note that if we ignore the vacuum state, which can not be sent by an ideal Alice, then $H^P = H^A$.

We use Corollary 1 to define attacks that cause no errors at all. For this specific implementation, \mathbf{U}_{zero} consists of the attacks that satisfy Equation (6) in four cases, matching the four BB84 states sent by Alice. Bob's setup (i.e. the constants $\beta_{k,j}$) is determined by the basis he measures; we denote the β 's imposed by the x -setup ($\phi = 0$) as $\beta_{k,j}^x$, and the ones imposed by the y -setup ($\phi = \pi/2$) by $\beta_{k,j}^y$, and write β in a matrix form. It is immediate from the operation of the interferometer Equation (7) that⁴

$$\beta_{k=\{t'_{-1}, t'_0, t'_1, t'_2\}, j=\{s_0, s_1, s_2, d_0, d_1, d_2\}}^x = \frac{1}{2} \begin{pmatrix} -1 & 0 & 0 & i & 0 & 0 \\ 1 & -1 & 0 & i & i & 0 \\ 0 & 1 & -1 & 0 & i & i \\ 0 & 0 & 1 & 0 & 0 & i \end{pmatrix}, \quad \beta_{k=\{t'_{-1}, t'_0, t'_1, t'_2\}, j=\{s_0, s_1, s_2, d_0, d_1, d_2\}}^y = \frac{1}{2} \begin{pmatrix} -i & 0 & 0 & -1 & 0 & 0 \\ 1 & -i & 0 & i & -1 & 0 \\ 0 & 1 & -i & 0 & i & -1 \\ 0 & 0 & 1 & 0 & 0 & i \end{pmatrix}. \quad (9)$$

Consider the case where Alice sends $|0_x\rangle$, namely, $\alpha_0 = \alpha_1 = \frac{1}{\sqrt{2}}$. An error occurs if Bob measures $|s_1\rangle$, $J_{\text{error}} = \{|s_1\rangle\}$, and by Equation (6), the attack causes no error if

$$-\frac{1}{2\sqrt{2}}(\epsilon_{0,0}|E_{0,0}\rangle_E + \epsilon_{1,0}|E_{1,0}\rangle_E) + \frac{1}{2\sqrt{2}}(\epsilon_{0,1}|E_{0,1}\rangle_E + \epsilon_{1,1}|E_{1,1}\rangle_E) = 0, \quad (10)$$

and when Alice sends $|1_x\rangle$, $J_{\text{error}} = \{|d_1\rangle\}$ which implies

$$\frac{i}{2\sqrt{2}}(\epsilon_{0,0}|E_{0,0}\rangle_E - \epsilon_{1,0}|E_{1,0}\rangle_E) + \frac{i}{2\sqrt{2}}(\epsilon_{0,1}|E_{0,1}\rangle_E - \epsilon_{1,1}|E_{1,1}\rangle_E) = 0. \quad (11)$$

³ Eve might have only partial control of these states, since they originate not only from the channel between Alice to Bob, but from the ancilla added by Bob as well.

⁴ While in this section we care only about $j = \{s_1, d_1\}$, in Section VI below we discuss $j = \{V, s_0, s_1, s_2, d_0, d_1, d_2\}$. Thus we describe here β 's elements for this entire set (omitting the vacuum state, since in any possible setup, $|V\rangle \rightarrow |V\rangle$).

The same is repeated for the y -basis and we get the solution

$$\epsilon_{0,0}|E_{0,0}\rangle_E = \epsilon_{1,1}|E_{1,1}\rangle_E \quad \text{and} \quad \epsilon_{1,0}|E_{1,0}\rangle_E = \epsilon_{0,1}|E_{0,1}\rangle_E = 0 . \quad (12)$$

It follows that Eve's attack must be of the form

$$\begin{aligned} |0\rangle_{\tilde{E}}|0_z\rangle_A &\xrightarrow{\mathcal{U}_E} p|\phi\rangle_E|0\rangle_P + \sqrt{(1-p)^2}|\psi_0\rangle_E|V\rangle_P \\ |0\rangle_{\tilde{E}}|1_z\rangle_A &\xrightarrow{\mathcal{U}_E} p|\phi\rangle_E|1\rangle_P + \sqrt{(1-p)^2}|\psi_1\rangle_E|V\rangle_P \end{aligned} \quad (13)$$

where p is the probability that Eve sends a qubit to Bob rather than blocking it. For $p < 1$ this is the *blocking* attack and for $p = 1$ it is the *identity* attack where Eve transmits the qubit to Bob intact. For these attacks, either Eve gets full information about Alice's state, or Bob does. In both cases, Eve has no information about the bits used for the key, and the specific implementation is *completely robust* according to Definition 1. \square

In Appendix C we generalize this result and prove robustness against a more realistic, yet limited, adversary restricted to pulses with at most two photons. Unfortunately, our method is not scalable to the general case.

V. TIME-BIN ENCODED STATES — “NATIVE” IMPLEMENTATION

A. “Native” implementation for x and z bases

Let us now extend the analysis to implementations that use the z -basis. This might be required, for instance, in order to implement the 6-state QKD protocol [47], in which Alice sends a qubit using the x , y and z bases at random; or in order to perform “QKD with classical Bob” [39, 48, 49] in which one party is restricted to use only the (classical) z -basis, and either performs measurements in that basis or returns the qubits (unchanged) to the other party.

We now describe a setup that Bob can employ in order to measure the z -basis, e.g. the states $|0_z\rangle = |t'_0\rangle$ and $|1_z\rangle = |t'_1\rangle$. The implementation is rather straightforward — Bob measures the pulses after the appropriate delay T_{short} , so that the measurement of $|0_z\rangle$ ($|1_z\rangle$) is done by opening a detector at time t_0 (t_1); see Figure 2.

The respective transformation \mathcal{U}_{B_z} is the identity operator

$$|0_z\rangle \xrightarrow{\mathcal{U}_{B_z}} |d_1\rangle \quad ; \quad |1_z\rangle \xrightarrow{\mathcal{U}_{B_z}} |s_1\rangle \quad (14)$$

where the other modes are not measured by Bob, and are not relevant for this scheme. We use the mode $|d_1\rangle$ instead of the more intuitive $|s_0\rangle$ in order to be consistent with the modes representing the bit values 0 and 1 when the x (or y) setup is used⁵.

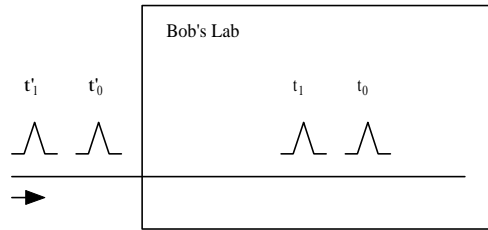


FIG. 2. Bob's laboratory setup for the z basis.

⁵ This can be justified by placing and removing a mirror, such that the pulse entering the lab at time t'_1 is reflected to the d arm.

We denote the BB84 protocol that uses x and z bases by alternating the setups (adding and removing beam-splitters as needed), as **native- xz -BB84**. In the same manner, a six-state protocol implemented by alternating the above setups (e.g., [50]) is denoted as **native-six-state** scheme.

B. Robustness of the native- xz -BB84 protocol against a single-photon-limited Eve

It is rather straightforward to extend the proof given in Section IV B to the native- xz -BB84 scheme. Define, for the z -basis setup, $J_0 = \{|d_1\rangle\}$, $J_1 = \{|s_1\rangle\}$ and $J_{\text{loss}} = I - J_0 - J_1$. The appropriate $\beta_{k,j}$, following Equation (14), are given by $\beta_{k=\{0,1\},j=\{s_1,d_1\}}^z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. It follows immediately from Equation (6) that any attack that causes no errors must satisfy $\epsilon_{0,1} = \epsilon_{1,0} = 0$. The set \mathbf{U}_{zero} includes attacks that satisfy the above requirement as well as the x -basis requirement. As before, the requirements yield the solution (13), and the native- xz -BB84 scheme is robust under our assumptions.

C. Robustness of the native-six-state protocol for a single-photon-limited Eve

It is easy to verify that the same result holds when using the y -basis instead of the x -basis. This proves the robustness of the native- yz -BB84 scheme. Combining this result with the result of the previous subsection immediately yields that the native-six-state scheme is robust as well, under the same assumptions.

Theorem 3. *Assuming a single-photon-limited adversary, the native-six-state scheme is completely robust.*

VI. TIME-BIN ENCODED STATES — “UNIFIED” IMPLEMENTATION

A. A “Unified” implementation for x and z bases

The native implementation suffers from one main caveat: the need of a mechanical operation after each qubit-transmission, as the basis must be chosen at random. Such an operation might take a lot of time and substantially decrease the maximal bit-rate allowed in the protocol. Other implementations do not involve mechanical operation but use a beam-splitter to split the channel such that each output reaches a different setup (see for instance [50]). These kinds of implementations suffer from a higher loss-rate and a lower bit-rate.

Let us describe a BB84 protocol that uses the z and x bases, in which Bob’s interferometric setup is fixed and independent of the basis used [20, 21, 23, 24]. The idea is to use the setup \mathcal{U}_{B_x} for measuring both bases in the following manner. In order to perform a measurement in the x basis, Bob opens his two detectors at time-bin t_1 , so that he measures the states $|s_1\rangle$ and $|d_1\rangle$. In addition, for measuring the z basis, Bob measures $|s_0\rangle, |d_0\rangle$ that implies the bit-value ‘0’ and $|s_2\rangle, |d_2\rangle$ that implies ‘1’, (see Equations (7) and (8)). We denote this scheme as **unified- xz -BB84**.

Bob measures different time-bins than t_1 , namely times t_0 and t_2 , and the set I becomes $\{|V\rangle, |d_0\rangle, |d_1\rangle, |d_2\rangle, |s_0\rangle, |s_1\rangle, |s_2\rangle\}$. In contrast to previous schemes, in this scheme the input modes t'_{-1} and t'_2 might have a non-zero overlap with the modes measured by Bob. The reverse-space-attack implies that the input space H^P is much larger than H^A : a state sent by Eve is a superposition of modes t'_{-1} to t'_2 .

B. (Non-)Robustness of unified- xz -BB84 scheme for a single-photon-limited Eve

Theorem 4. *The unified- xz -BB84 scheme is completely nonrobust*

Proof. Let us repeat the above analysis for the unified- xz -BB84 scheme. For the non-robustness proof, it suffices to restrict Eve (as well as any natural noise) to single photon pulses. The requirements for the x -basis remain as given in Section IV B (Equations (10)–(11)). In addition, when Bob measures the z basis, he interprets his outcome according to $J_0 = \{|d_0\rangle, |s_0\rangle\}$, $J_1 = \{|d_2\rangle, |s_2\rangle\}$, $J_{\text{invalid}} = \{\}$ (due to the single-photon assumption), and $J_{\text{loss}} = I - J_0 - J_1$. The setup is joined for both the z and the x bases, $\beta^z = \beta^x$, whose value is given in Equation (9).

Following Corollary 1, an attack \mathcal{U}_E causes no errors if it satisfies

$$i\epsilon_{0,1}|E_{0,1}\rangle + i\epsilon_{0,2}|E_{0,2}\rangle = 0 \quad -\epsilon_{0,1}|E_{0,1}\rangle + \epsilon_{0,2}|E_{0,2}\rangle = 0 \quad (15)$$

corresponding to the case where Alice sends $|0_z\rangle$, i.e. $\alpha_0 = 1$, $\alpha_1 = 0$, and $J_{\text{error}} = \{|d_2\rangle, |s_2\rangle\}$, as well as

$$i\epsilon_{1,-1}|E_{1,-1}\rangle + i\epsilon_{1,0}|E_{1,0}\rangle = 0 \quad -\epsilon_{1,-1}|E_{1,-1}\rangle + \epsilon_{1,0}|E_{1,0}\rangle = 0 \quad (16)$$

corresponding to the case where Alice sends $|1_z\rangle$, i.e. $\alpha_0 = 0$, $\alpha_1 = 1$, and $J_{\text{error}} = \{|d_0\rangle, |s_0\rangle\}$. This leads to the constraints $\epsilon_{0,1} = \epsilon_{0,2} = 0$ and $\epsilon_{1,-1} = \epsilon_{1,0} = 0$. Along with the requirements for the x -basis the only possible attacks are of the form

$$\begin{aligned} |0\rangle_{\tilde{E}}|0_z\rangle_A &\xrightarrow{\mathcal{U}_E} p|\phi\rangle_E|0\rangle_P + p_1|\phi_1\rangle|-1\rangle_P + p_2|\psi_0\rangle_E|V\rangle_P \\ |0\rangle_{\tilde{E}}|1_z\rangle_A &\xrightarrow{\mathcal{U}_E} p|\phi\rangle_E|1\rangle_P + p_3|\phi_2\rangle|2\rangle_P + p_4|\psi_1\rangle_E|V\rangle_P \end{aligned} \quad (17)$$

with $|p|^2 + |p_1|^2 + |p_2|^2 = |p|^2 + |p_3|^2 + |p_4|^2 = 1$. Using this result, it is easy to devise an attack and show the protocol is completely non-robust. For instance, let

$$|0\rangle_{\tilde{E}}|0_z\rangle_A \xrightarrow{\mathcal{U}_E} |E_1\rangle_E|t'_{-1}\rangle_P \quad |0\rangle_{\tilde{E}}|1_z\rangle_A \xrightarrow{\mathcal{U}_E} |E_2\rangle_E|t'_2\rangle_P$$

with orthogonal $|E_1\rangle$, $|E_2\rangle$. This attack never causes an error, yet it increases the loss rate—Bob always gets a loss when using the x basis. This means that only bits encoded using the z basis are used for transferring information, and Eve can copy the information. It follows that the unified- xz -BB84 is *completely nonrobust* according to Definition 1. (This specific attack is somewhat related to the “fake state” attack of [34]). \square

As mentioned, in the above attack all the qubits passed by Eve are in the z -basis (i.e., the loss-rate Bob sees for the x -basis is 1). We can compose an attack that doesn’t have such a property (for instance, by letting $p > 0$), in which Eve does not force a loss in the x -basis, yet she does not learn the information for that basis.

Finally, note that the above attack also applies to the unified six-state QKD scheme (see [22], for instance), making such realizations totally insecure. Going beyond the schemes presented here is left for future research.

VII. CONCLUSION

We study robustness of common QKD implementations by formulating the conditions that make a specific attack undetectable. Assuming a single-photon restricted adversary, we show that several

of the implementations in use are robust, while others are completely nonrobust (and thus insecure): the adversary is capable of learning all of the information about the final key without causing any errors at the legitimate user's end. The security flaw emerges from the way the devices are used, rather than their imperfections or insecurity of the underlying BB84 (or six-state) protocol. A complete security proof of the implementations we prove robust is still missing. Another question we leave open is whether the above robustness proof can be extended to the case of an unlimited adversary.

We conclude that a security analysis of a QKD realization must be done according to the specific equipment in use. A security proof of a theoretical protocol is relevant only when the setup considered is proven to realize the theoretical protocol in an exact manner. Yet, any realization deviates from the theoretical one, and this should be considered by the security analysis. A general framework for all possible deviations is still missing (see discussion in a preliminary version of this paper [51]). This conclusion is crucial when considering off-the-shelf products, claiming to bring QKD with proven security.

ACKNOWLEDGMENTS

The authors would like to thank Michel Boyer for many useful suggestions. We also wish to thank Akshay Wadia, Alan Roytman and Niek Bouman for miscellaneous comments.

-
- [1] C. A. Fuchs and A. Peres, *Physical Review A* **53**, 2038 (1996).
 - [2] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, *Physical Review A* **56**, 1163 (1997).
 - [3] C. A. Fuchs, *Fortschritte der Physik* **46**, 535 (1999).
 - [4] E. Biham, M. Boyer, G. Brassard, J. van de Graaf, and T. Mor, *Algorithmica* **34**, 372 (2002).
 - [5] P. O. Boykin and V. P. Roychowdhury, "Information vs. disturbance in dimension D," (2004), [quant-ph/0412028](https://arxiv.org/abs/quant-ph/0412028).
 - [6] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. P. Roychowdhury, *J. Cryptology* **19**, 381 (2006).
 - [7] C. H. Bennett and G. Brassard, *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, 175 (1984).
 - [8] C. H. Bennett, *Physical Review Letters* **68**, 3121 (1992).
 - [9] P. D. Townsend, *Electronics Letters* **30**, 809 (1994).
 - [10] R. J. Hughes, G. G. Luther, G. L. Morgan, and C. Simmons, in *Rochester conference on coherence and quantum optics* (Rochester, NY (United States), 1995) pp. 7–10.
 - [11] C. Gobby, Z. L. Yuan, and A. J. Shields, *Applied Physics Letters* **84**, 3762 (2004).
 - [12] K. Inoue, E. Waks, and Y. Yamamoto, *Physical Review Letters* **89**, 037902 (2002).
 - [13] H. Takesue, E. Diamanti, T. Honjo, C. Langrock, M. M. Fejer, K. Inoue, and Y. Yamamoto, *New Journal of Physics* **7**, 232 (2005).
 - [14] H. Takesue, T. Honjo, and H. Kamada, *Japanese Journal of Applied Physics* **45**, 5757 (2006).
 - [15] A. Muller, T. Herzog, B. Huttner, W. Tittel, H. Zbinden, and N. Gisin, *Applied Physics Letters* **70**, 793 (1997).
 - [16] D. Stucki, N. Gisin, O. Guinnard, G. Ribordy, and H. Zbinden, *New Journal of Physics* **4**, 41 (2002).
 - [17] <http://www.idquantique.com/>.
 - [18] <http://www.magiqtech.com/>.
 - [19] G. Bonfrate, M. Harlow, C. Ford, G. Maxwell, and P. Townsend, *Electronics Letters* **37**, 846 (2001).
 - [20] Y. Nambu, T. Hatanaka, and K. Nakamura, "Planar lightwave circuits for quantum cryptographic systems," (2003).
 - [21] Y. Nambu, T. Hatanaka, and K. Nakamura, *Japanese Journal of Applied Physics* **43**, L1109 (2004).
 - [22] M. Nazarathy, I. Tselniker, Y. Regev, M. Orenstein, and M. Katz, *Selected Topics in Quantum Electronics, IEEE Journal of* **12**, 897 (2006).

- [23] Z. D. Walton, A. F. Abouraddy, A. V. Sergienko, B. E. A. Saleh, and M. C. Teich, *Phys. Rev. Lett.* **91**, 087901 (2003).
- [24] G. Jaeger and A. Sergienko, *AIP Conference Proceedings* **810**, 161 (2006).
- [25] M. Gao, L.-M. Liang, C.-Z. Li, and C.-L. Tian, *Physics Letters A* **359**, 126 (2006).
- [26] P. W. Shor and J. Preskill, *Physical Review Letters* **85**, 441 (2000).
- [27] D. Mayers, *J. ACM* **48**, 351 (2001).
- [28] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, Swiss Federal Institute Of Technology, Zurich (2005), [quant-ph/0512258](#).
- [29] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, in *TCC* (Springer, 2005) pp. 386–406.
- [30] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders, *Physical Review Letters* **85**, 1330 (2000).
- [31] N. Lütkenhaus, *Physical Review A* **61**, 052304 (2000).
- [32] E. Waks, H. Takesue, and Y. Yamamoto, *Physical Review A (Atomic, Molecular, and Optical Physics)* **73**, 012344 (2006).
- [33] D. Mayers and A. Yao, in *FOCS '98: Proceedings of the 39th Annual Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, USA, 1998) p. 503.
- [34] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, *Quantum Information and Computation* **5**, 325 (2004).
- [35] H. Inamori, N. Lütkenhaus, and D. Mayers, *European Physical Journal D* **41**, 599 (2007).
- [36] E. Davies, *Information Theory*, *IEEE Transactions on* **24**, 596 (1978).
- [37] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, *Physical Review Letters* **92**, 057901 (2004).
- [38] C. H. Bennett, G. Brassard, and N. D. Mermin, *Physical Review Letters* **68**, 557 (1992).
- [39] M. Boyer, D. Kenigsberg, and T. Mor, *Physical Review Letters* **99**, 140501 (2007).
- [40] V. Makarov and D. R. Hjelm, *Journal of Modern Optics* **52**, 691 (2005).
- [41] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, *Phys. Rev.* **134**, B1410 (1964).
- [42] Y. Aharonov and L. Vaidman, *Phys. Rev. A* **41**, 11 (1990).
- [43] E. Biham, B. Huttner, and T. Mor, *Phys. Rev. A* **54**, 2651 (1996).
- [44] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Reviews of Modern Physics* **74**, 145 (2002).
- [45] C. Elliott, D. Pearson, and G. Troxel, in *SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications* (ACM Press, New York, NY, USA, 2003) pp. 227–238.
- [46] M. Dusek, N. Lütkenhaus, and M. Hendrych, in *Progress in Optics*, Vol. 49, edited by E. Wolf (Elsevier, 2006) pp. 381 – 454.
- [47] D. Bruß, *Physical Review Letters* **81**, 3018 (1998).
- [48] M. Boyer, R. Gelles, D. Kenigsberg, and T. Mor, *Phys. Rev. A* **79**, 032341 (2009).
- [49] X. Zou, D. Qiu, L. Li, L. Wu, and L. Li, *Phys. Rev. A* **79**, 052312 (2009).
- [50] M. Nazarathy, *Opt. Lett.* **30**, 1533 (2005).
- [51] R. Gelles and T. Mor, “Quantum-Space Attacks,” (2007), [arXiv:0711.3019](#).
- [52] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Physical Review Letters* **73**, 58 (1994).
- [53] C. Gerry and P. Knight, *Introductory Quantum Optics* (Introductory Quantum Optics, by Christopher Gerry and Peter Knight, pp. 332. ISBN 0521820359. Cambridge, UK: Cambridge University Press, November 2004., 2004).
- [54] N. Lütkenhaus, J. Calsamiglia, and K.-A. Suominen, *Physical Review A* **59**, 3295 (1999).
- [55] E. Knill, R. Laflamme, and G. J. Milburn, *Nature (London)* **409**, 46 (2001).
- [56] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, *Reviews of Modern Physics* **79**, 135 (2007).
- [57] J. Kim, S. Takeuchi, Y. Yamamoto, and H. H. Hogue, *Applied Physics Letters* **74**, 902 (1999).
- [58] D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, and I. A. Walmsley, *Opt. Lett.* **28**, 2387 (2003).
- [59] O. Haderka, M. Hamar, and J. Peřina, *European Physical Journal D* **28**, 149 (2004).

APPENDIX

Appendix A: Photonic Qubits and Fock Space

The Fock-Space (FS) notation is the best way to describe a quantum system where the “players” are indistinguishable particles such as photons, using the occupancy number basis. The Fock-state $|n\rangle^F$ represents n particles in a given mode⁶, for instance, the number of photons in a certain electromagnetic pulse that have the same horizontal polarization $|\leftrightarrow\rangle$. When needed, a subscript is added to the Fock-state in order to identify the specific mode, e.g. $|n\rangle_{\uparrow}^F$ or $|m\rangle_{\leftrightarrow}^F$. When more than one mode is considered, we write the joint state $|n_1, n_2, \dots, n_k\rangle^F$ to indicate n_i photons in the i th mode. Using this notation, a description of a general single-photon qubit, $|\phi_{\text{qubit}}\rangle = \alpha_0|10\rangle^F + \alpha_1|01\rangle^F$, is based on using two modes (say two orthogonal polarizations) $|0_z\rangle \equiv |10\rangle^F$ and $|1_z\rangle \equiv |01\rangle^F$. For instance, the states $|0_x\rangle \equiv \frac{1}{\sqrt{2}}(|10\rangle^F + |01\rangle^F)$ and $|1_x\rangle \equiv \frac{1}{\sqrt{2}}(|10\rangle^F - |01\rangle^F)$ commonly represent the two diagonal polarizations.

Unfortunately, in real life Alice is unable to send perfect qubits; due to the specific device used, Alice often sends the vacuum state $|00\rangle^F$, and also sometimes sends more than a single photon (i.e., the states $|20\rangle^F, |11\rangle^F$ and $|02\rangle^F$). To be more precise, she actually sends the 2-mode multi-photon state $\sum_{n_1=0, n_2=0}^{\infty} \alpha_{n_1, n_2} |n_1, n_2\rangle^F$, containing also terms with more than two photons. Such terms usually have a negligible probability, and it is sufficient to analyze the 6-dimensional Hilbert space of zero, one and two photons. Alice might also (unintentionally) send more modes than she intended to. Thus, the most general state Alice could send is a k -mode multi-photon state $\sum_{n_1, \dots, n_k=0}^{\infty} \alpha_{n_1, \dots, n_k} |n_1, \dots, n_k\rangle^F$. Sending more than two modes could also have a negative effect on the security of the protocol.

Bob’s ideal measurement of the Fock-state $|n\rangle^F$ is commonly assumed to be limited to a complete measurement that yields the number of photons occupying the mode, i.e. the number n . This can be extended to an ideal measurement of the k -mode Fock-state $|n_1, n_2, \dots, n_k\rangle^F$ which yields the numbers n_1 to n_k . In Appendix C 1, we discuss more realistic measurements of a multi-photon state.

In addition, Bob can measure other specific properties of the state using (for instance) beam splitters, phase shifters and mirrors [52]. For example, let us assume that Bob wants to distinguish the state $\frac{1}{\sqrt{2}}(|10\rangle^F + |01\rangle^F)$ from $\frac{1}{\sqrt{2}}(|10\rangle^F - |01\rangle^F)$, where the different modes are different paths of the photon. Bob can perform a phase shift of 45° on the path represented by the first mode, and then place a symmetric beam splitter to obtain $|10\rangle^F$ or $|01\rangle^F$ respectively at the outputs of the beam splitter (up to a general phase). These two states can be distinguished by a simple measurement as described above.

Appendix B: Interferometer

An interferometer (Figure 1) is a device composed of two beam splitters (BS) with one short path, one long path, and a controlled phase shifter P_ϕ , that is placed at the long arm of the interferometer. We focus on the following case which is used for measuring differential phase-shift QKD.

In each transmission, a superposition of two (time) modes enter the interferometer and result in a superposition of 6 modes (Figure 1). The input modes are separated with a time difference of ΔT seconds, that is, the first mode arrives at time t'_0 , and the second at $t'_1 = t'_0 + \Delta T$. The first pulse travels through the short arm in T_{short} seconds, and through the long arm in $T_{\text{long}} = T_{\text{short}} + \Delta T$

⁶ We use the notation $|\cdot\rangle^F$ to indicate use of the occupancy number basis.

seconds, where the time difference between the two arms is exactly the time difference ΔT between the two incoming modes. Due to traveling through both arms, the first mode yields outgoing pulses both at time $t_0 \equiv t'_0 + T_{short}$ and at $t_1 \equiv t'_0 + T_{long} = t'_0 + T_{short} + \Delta T = t_0 + \Delta T$.

When the second pulse enters the interferometer, it also travels through both arms. Intuitively, the part of the t'_1 mode that travels through the short arm interferes with the part of the t'_0 mode that travels through the long arm, and the output exits the interferometer at t_1 . The part of the second pulse that travels through the long arm exits the interferometer at time $t_2 = t_1 + \Delta T$. As a result, we can actually see six pulses at the two output arms, three in each direction, with the two middle pulses determined by the interference between the two pulses arriving into Bob's lab. We shall now write this formally.

1. Beam splitter

Each one of the beam splitters has two input arms (modes 1, 2) and two output arms (modes 3, 4), see Figure 3. Each entering photon is transmitted (or reflected) with probability 0.5; The transmitted part keeps the same phase as the incoming photon, while the reflected part gets an extra phase of $e^{i\pi/2}$. Specifically, $|10\rangle_{1,2}^F \rightarrow \frac{1}{\sqrt{2}}(|10\rangle_{3,4}^F + i|01\rangle_{3,4}^F)$ and $|01\rangle_{1,2}^F \rightarrow \frac{1}{\sqrt{2}}(i|10\rangle_{3,4}^F + |01\rangle_{3,4}^F)$. Thus, for a single photon state, the transformation is of the form

$$\alpha|10\rangle_{1,2}^F + \beta|01\rangle_{1,2}^F \mapsto \frac{\alpha + i\beta}{\sqrt{2}}|10\rangle_{3,4}^F + \frac{i\alpha + \beta}{\sqrt{2}}|01\rangle_{3,4}^F. \quad (\text{B1})$$

It is important to note that when a single mode (carrying a single photon) enters a beam splitter from one arm, and nothing (namely, vacuum) enters the other arm (say, $\alpha = 1; \beta = 0$), there are still *two* output modes. This means that the other (vacuum) entry must be considered as an additional mode — an ancilla carrying no photons.

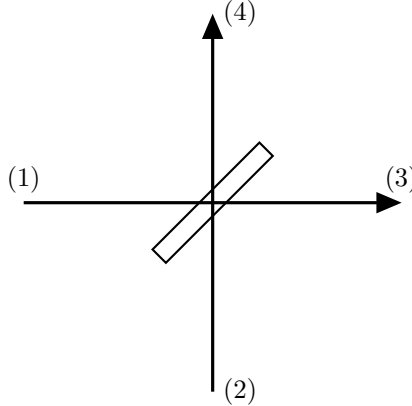


FIG. 3. A symmetric beam-splitter with two input modes (1) and (2) and two output modes (3) and (4).

2. Phase shifter

The controlled phase shifter P_ϕ performs a phase shift on the input state by a given phase ϕ , i.e. $P_\phi(|n\rangle^F) = e^{i n \cdot \phi} |n\rangle^F$, see [53]. The users can change the phase according to the specific basis in use. Clearly, the transformation changes only the mode which travels through the phase shifter (on the long arm), while the other modes do not change.

3. Evolution of a single pulse through the interferometer

When a single mode, carrying one or more photons, enters the interferometer, three ancillas in a vacuum state are added by the interferometric setup (see Figure 4). As mentioned above, the mode entering the interferometer at time t'_0 , yields two modes at time t_0 , and two modes at time t_1 . These four output modes are: times t_0, t_1 at the 's' (straight) arm of the interferometer, and times t_0, t_1 at the 'd' (down) arm of the interferometer. A basis state of this Fock-space can be written as $|n_{s0}, n_{s1}, n_{d0}, n_{d1}\rangle^F$.

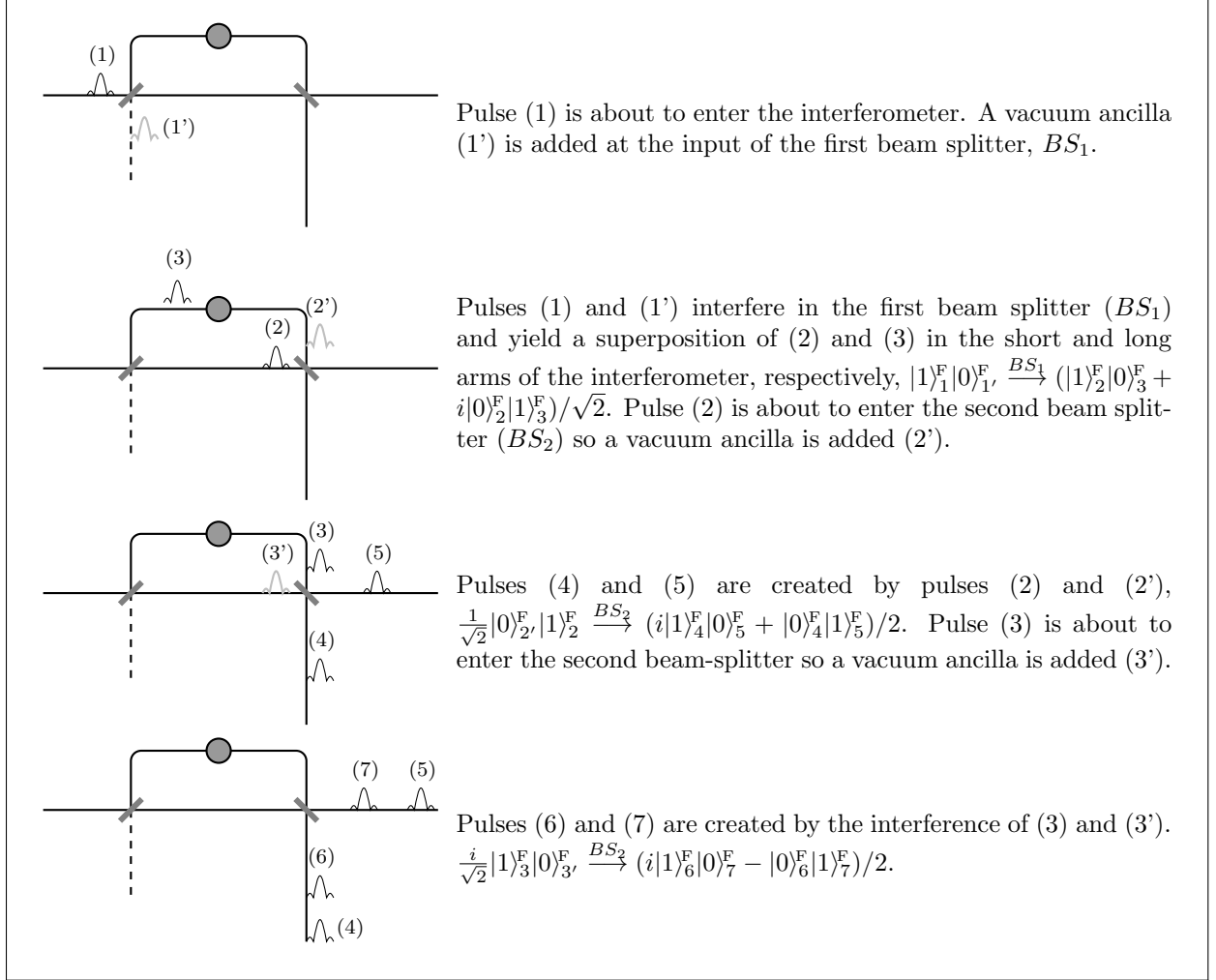


FIG. 4. Evolution in time of a single photon pulse through the interferometer with $\phi = 0$, $|1000\rangle_{1,1',2',3'}^F \rightarrow \frac{1}{2}(|1000\rangle^F - |0100\rangle^F + i|0010\rangle^F + i|0001\rangle^F)_{5,7,4,6}$. The output state is denoted by modes $|n_{s0}, n_{s1}, n_{d0}, n_{d1}\rangle^F$ that correspond to modes (5), (7), (4) and (6) respectively.

Assume that a single photon enters the interferometer at time t'_0 . Using the above notations, the interferometer's transformation is given by

$$|1\rangle_{t'_0}^F |000\rangle^F \mapsto (|1000\rangle^F - e^{i\phi}|0100\rangle^F + i|0010\rangle^F + ie^{i\phi}|0001\rangle^F) / 2. \quad (B2)$$

Note the three vacuum ancillas that were added. Also note that a pulse which is sent at a different time (say, t'_1 , or t'_{-1} , etc.) results in the same output state, with appropriate delays. That is, a

pulse entering the interferometer at time t'_i results in the state $(|1000\rangle^F - e^{i\phi}|0100\rangle^F + i|0010\rangle^F + ie^{i\phi}|0001\rangle^F) / 2$ in a Fock-space with basis states $|n_{s_i}, n_{s_{i+1}}, n_{d_i}, n_{d_{i+1}}\rangle^F$.

4. Evolution of two pulses through the interferometer

We are now ready to consider the setup of Figure 1 and two input modes, t'_0 and t'_1 , that enter the interferometer one after the other, with exactly the same time difference ΔT as the interferometer's arms. As a result of this precise timing, the two modes are transformed into a superposition of only six modes (instead of eight modes) at the outputs (see Figure 5). Four (vacuum state) ancillas are added during the process and the resulting six modes are t_0, t_1, t_2 at the 's' arm and the 'd' arm of the interferometer. A basis state of this Fock-space is therefore $|n_{s_0}, n_{s_1}, n_{s_2}, n_{d_0}, n_{d_1}, n_{d_2}\rangle^F$. If exactly one photon enters the interferometer, we can use Equation (B2) to obtain

$$\begin{aligned} |1\rangle_{t'_0}^F |0\rangle_{t'_1}^F |0000\rangle^F &\mapsto (|100000\rangle^F - e^{i\phi}|010000\rangle^F + i|000100\rangle^F + ie^{i\phi}|000010\rangle^F) / 2 \\ |0\rangle_{t'_0}^F |1\rangle_{t'_1}^F |0000\rangle^F &\mapsto (|010000\rangle^F - e^{i\phi}|001000\rangle^F + i|000010\rangle^F + ie^{i\phi}|000001\rangle^F) / 2 \end{aligned} \quad (\text{B3})$$

Recall that $|0_z\rangle = |10\rangle_{t'_0 t'_1}^F$ and $|1_z\rangle = |01\rangle_{t'_0 t'_1}^F$. It follows that an arbitrary qubit is transformed as

$$\begin{aligned} &(\alpha|10\rangle^F + \beta|01\rangle^F) |0000\rangle^F \longrightarrow \\ &\left(\frac{\alpha}{2} |100000\rangle^F + \frac{\beta - \alpha e^{i\phi}}{2} |010000\rangle^F - \frac{\beta e^{i\phi}}{2} |001000\rangle^F + \frac{i\alpha}{2} |000100\rangle^F + \frac{i(\alpha e^{i\phi} + \beta)}{2} |000010\rangle^F + \frac{i\beta e^{i\phi}}{2} |000001\rangle^F \right). \end{aligned} \quad (\text{B4})$$

Appendix C: Robustness of *xy*-BB84 against a more realistic Eve

In this section we prove that the *xy*-BB84 scheme is robust against an adversary that can send pulses with up to 2 photons. Although we believe that the protocol is robust against an unlimited Eve our proof is not scalable to the general case.

1. Measurement of More Than One Photon

Recall that a general (ideal) measurement of the k -mode Fock-state $|n_1, n_2, \dots, n_k\rangle^F$ yields the numbers n_1 to n_k . However, we assume that Bob uses imperfect devices which might restrict him to perform only limited measurement [54–56]. A realistic Bob performs an incomplete measurement, in which some modes might not be measured, and some modes he cannot detect the exact number of photons. For instance, Bob might measure a mode i using a threshold detector and only determine whether the mode is empty or non-empty (i.e., whether or not the number of received photons equals 0), described as the projection $P_i = \sum_{n_i=1}^{\infty} |n_1, n_2, \dots, n_k\rangle^F \langle n_1, n_2, \dots, n_k|$, where $n_j = 0$ for $j \neq i$. A better measurement could (theoretically) allow him to distinguish the exact number of photons populating the mode; this is done using a device named a *counter*, also known as a *photon-number-resolving* detector [57–59].

In the following, we consider the case in which Bob's detectors cannot distinguish between detecting a single photon or more. We stress that our robustness proof holds even for the limited (realistic) Bob described above, and not only for an ideal Bob.

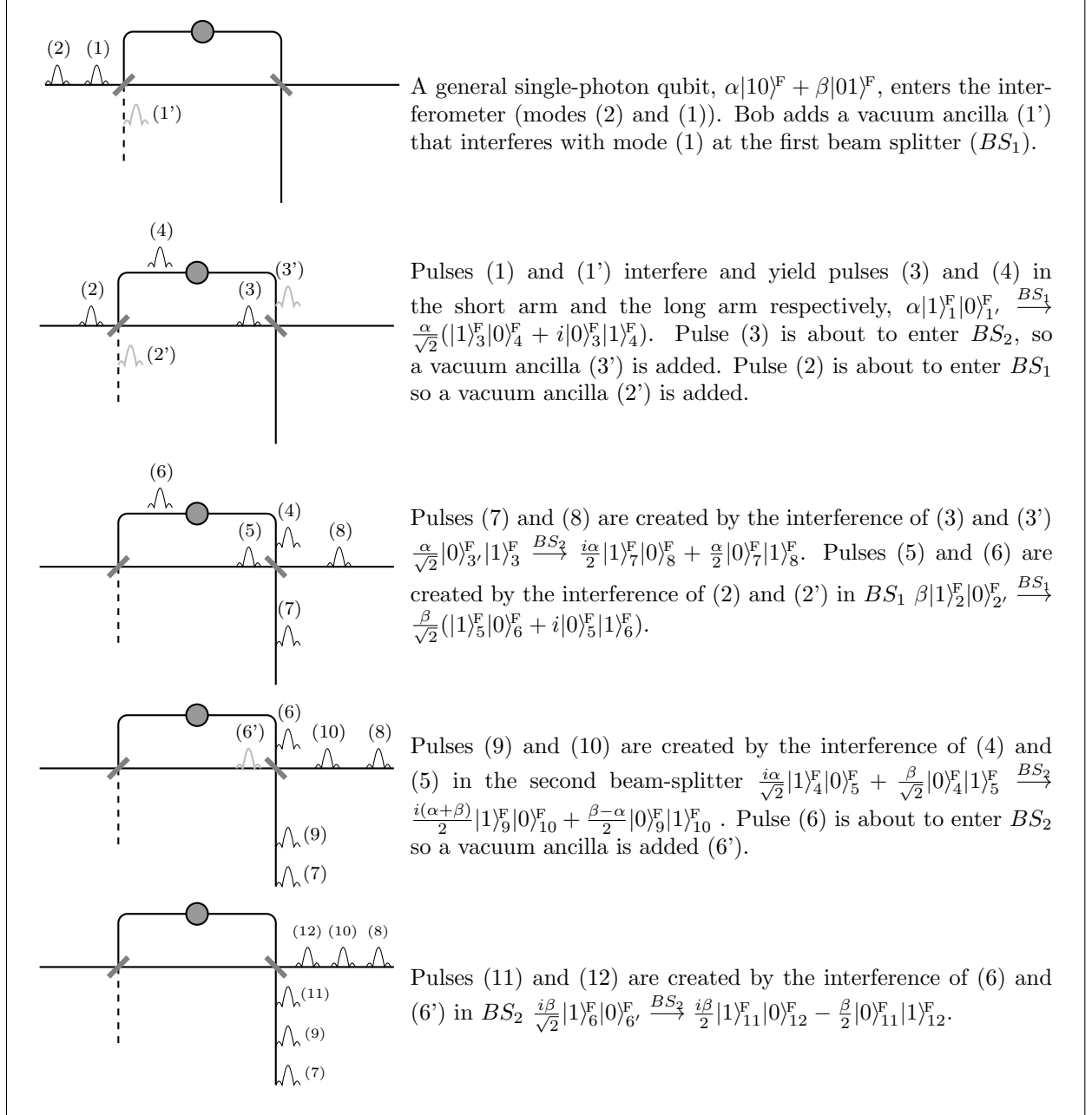


FIG. 5. Evolution in time of two modes through the interferometer with $\phi = 0$, $(\alpha|1\rangle_1^F|0\rangle_2^F + \beta|0\rangle_1^F|1\rangle_2^F)|0000\rangle_{1',2',3',6'}^F \rightarrow (\frac{\alpha}{2}|100000\rangle^F + \frac{\beta-\alpha}{2}|010000\rangle^F - \frac{\beta}{2}|001000\rangle^F + \frac{i\alpha}{2}|000100\rangle^F + \frac{i(\alpha+\beta)}{2}|000010\rangle^F + \frac{i\beta}{2}|000001\rangle^F)_{8,10,12,7,9,11}^F$. The output state is denoted by modes $|n_{s_0}, n_{s_1}, n_{s_2}, n_{d_0}, n_{d_1}, n_{d_2}\rangle^F$.

2. A Robustness Proof for the xy -BB84 Scheme

We begin by extending the interferometer transformation for pulses of two photons.

Proposition 5. *If the time modes t'_0 and t'_1 contain exactly 2 photons, the pulse evolves in the*

interferometer in the following manner:

$$U_{IT}|2\rangle_{t'_0}^F = \frac{1}{4} \left[|200000\rangle^F + \sqrt{2}i|100100\rangle^F - |000200\rangle^F \right. \\ \left. + e^{i\phi}(-\sqrt{2}|110000\rangle^F + \sqrt{2}i|100010\rangle^F - \sqrt{2}i|010100\rangle^F - \sqrt{2}|000110\rangle^F) \right. \\ \left. + e^{i \cdot 2\phi}(|020000\rangle^F - \sqrt{2}i|010010\rangle^F - |000020\rangle^F) \right] \quad (C1)$$

$$U_{IT}|11\rangle_{t'_0 t'_1}^F = \frac{1}{4} \left[|110000\rangle^F + i|100010\rangle^F + i|010100\rangle^F - |000110\rangle^F \right. \\ \left. + e^{i\phi}(-\sqrt{2}|020000\rangle^F - \sqrt{2}|000020\rangle^F - |101000\rangle^F \right. \\ \left. + i|100001\rangle^F - i|001100\rangle^F - |000101\rangle^F) \right. \\ \left. + e^{i \cdot 2\phi}(|011000\rangle^F - i|010001\rangle^F - i|001010\rangle^F - |000011\rangle^F) \right] \quad (C2)$$

with basis $|n_{s_0}, n_{s_1}, n_{s_2}, n_{d_0}, n_{d_1}, n_{d_2}\rangle^F$. Note that $U_{IT}|2\rangle_{t'_1}^F$ is immediate from $U_{IT}|2\rangle_{t'_0}^F$.

Proof. Consider the evolution of a pulse with 2 photons through a beam-splitter [53]

$$|20\rangle_{1,2}^F \xrightarrow{BS} \frac{1}{2}(|20\rangle^F + \sqrt{2}i|11\rangle^F - |02\rangle^F)_{3,4} \\ |02\rangle_{1,2}^F \xrightarrow{BS} \frac{1}{2}(-|20\rangle^F + \sqrt{2}i|11\rangle^F + |02\rangle^F)_{3,4} \\ |11\rangle_{1,2}^F \xrightarrow{BS} \frac{i}{\sqrt{2}}(|20\rangle^F + |02\rangle^F)_{3,4}$$

Composing the two beam-splitters and phase shift, with the appropriate delays immediately leads to the above result. \square

Theorem 6. *The xy -BB84 scheme is robust against an attack limited to pulses with at most two photons.*

Proof. We consider three different cases, according to the number of photons occupying the modes t'_0 and t'_1 :

1. There are 0 photons in modes t'_0 and t'_1 : For instance, Eve sends the state $|11\rangle_{t'_4 t'_5}^F$ or $|20\rangle_{t'_{-9} t'_6}^F$. These kinds of states always cause a loss since they never reach Bob's detectors at time t_1 . Eve can send a superposition of any such states, and Bob will not be able to distinguish them from the case that Eve sends $|V\rangle$. Such states are in J_{loss} and do not affect the error rate.
2. There is a single photon in modes t'_0 and t'_1 : For instance, Eve sends $|11\rangle_{t'_0 t'_3}^F$. This case is not the same as sending a single photon, but since only one photon can be measured at time t_1 , we can consider this case as giving Eve the second photon, instead of sending it to Bob. Thus, we can use the robust proof for a single photon of Section IV B to deal with this case as well.
3. Both photons are in modes t'_0 and t'_1 : Assume Eve sends Bob a 2-photon state of the form $|\psi_{\alpha, \beta, \gamma}\rangle = \alpha|20\rangle_{t'_0 t'_1}^F + \beta|11\rangle_{t'_0 t'_1}^F + \gamma|02\rangle_{t'_0 t'_1}^F$. We show that if Eve causes no errors, Corollary 1 restricts the state to be a multiple-photon version of the state expected by Bob. E.g. if Bob expects $|0_x\rangle$, the only 2-photon state that causes no errors is $|0_x\rangle^{(2)} = \frac{1}{2}(|20\rangle_{t'_0 t'_1}^F + \sqrt{2}|11\rangle_{t'_0 t'_1}^F + |02\rangle_{t'_0 t'_1}^F)$.

We demonstrate the above only for the case of $|0_x\rangle$. The other cases can be achieved in a similar way. Recall that in this case an error happens if Bob's d_1 detector clicks. This means that in order to have no errors, we must require that ${}^F\langle abcdef|U_{IT}|\psi_{\alpha,\beta,\gamma}\rangle = 0$ for any element with $b > 0$.

A zero overlap with $|020000\rangle^F$ implies $\alpha - \sqrt{2}\beta + \gamma = 0$. The element $|010010\rangle^F \in J_{\text{invalid}}$ since it causes a click in both detectors, which implies an inconclusive result (that must be interpreted as an error, since Alice's states are assumed to contain only a single photon). Since ${}^F\langle 010010|U_{IT}|20\rangle_{t'_0 t'_1}^F = -{}^F\langle 010010|U_{IT}|02\rangle_{t'_0 t'_1}^F$ and ${}^F\langle 010010|U_{IT}|11\rangle_{t'_0 t'_1}^F = 0$ a zero-overlap requires $\alpha = \gamma$. The only solution for the above constraints is $\beta = \sqrt{2}\alpha = \sqrt{2}\gamma$, and one can easily verify that $\frac{1}{2}|\psi_{\alpha=1,\beta=\sqrt{2},\gamma=1}\rangle$ causes no error if Bob expects $|0_x\rangle$, i.e. Bob's detector at s_1 will never click.

Note that Eve cannot perform an attack that always sends Bob a perfect 2-photon copy of the state sent by Alice, due to no-cloning.

Last, we consider superpositions of states with different numbers of photons (e.g., $(|01\rangle_{t'_0 t'_1}^F + |20\rangle_{t'_0 t'_1}^F)/\sqrt{2}$, etc.). We note that the overlap of states with different numbers of photons is always zero. Since $\langle m_1, m_2, \dots, m_k | m'_1, m'_2, \dots, m'_k \rangle^F = \prod_{i=\{1..k\}} \delta_{m_i m'_i}$, then if $\sum_i m \neq \sum_i m'$, the overlap must be zero. It follows that in order to have zero overlap with a state $|\phi\rangle \in J_{\text{error}} \cup J_{\text{invalid}}$ we only need to consider pulses with the same number of photons as $|\phi\rangle$. \square